Fertility and Life Expectancy: Predictive (A)symmetries across OECD Countries, 2000-2024

A Unified Perspective on Formation and Dissolution Processes in Demography
,
The Interdisciplinary Centre on Population Dynamics (CPop), University of Southern Denmark, 9-11 March,2026

LINKÖPING UNIVERSITY

Miloš Milovanović
Divison of Ageing and Social Change (ASC), Department of Culture and Society (IKOS)

Introduction

Fertility and longevity are jointly embedded in institutional and socio-demographic context.

Examining them within an identical analytical framework may help clarify the extent to which these processes reflect common macro-demographic regimes.

Aim

This study examines whether fertility and life expectancy exhibit symmetric or asymmetric predictive structures when modelled as outcomes of the same sociodemographic conditions.

The analysis seeks to explore how both outcomes relate to identical contextual conditions across countries and over time.

Data

The study is based on annual data from 31 OECD member countries included in the OECD database.

The analysis excluded six countries (Chile, Colombia, Mexico, Türkiye, Costa Rica, and the United States), because the full data set for 2000–2024 was not available.

Israel is excluded due to the much higher values of TFR.

Outcomes and Predictors

Outcomes-Total Fertility Rate (TFR) and Life Expectancy at Birth (LE)

edu- share of woman with university degree.

socpast- “1” country was member of the socialist block, otherwise “0”.

fampatt-We grouped countries into clusters based on policy configurations / welfare regime types that share similar social and family policies, following Oláh, Kotowska, and Richter (2018).

region- The region is a variable operationalized by clustering countries into five groups based on their similarity in historical development and geographical proximity.

socspen- it includes government spending on: Old-age pensions, Health-related benefits, Family and child benefits, Disability benefits, Unemployment benefits, Housing support.

emplo and emplopar- The female employment rate and part-time employment rate.

gwg- The gender wage gap is defined as the difference between the median earnings of women relative to the median earnings of men.

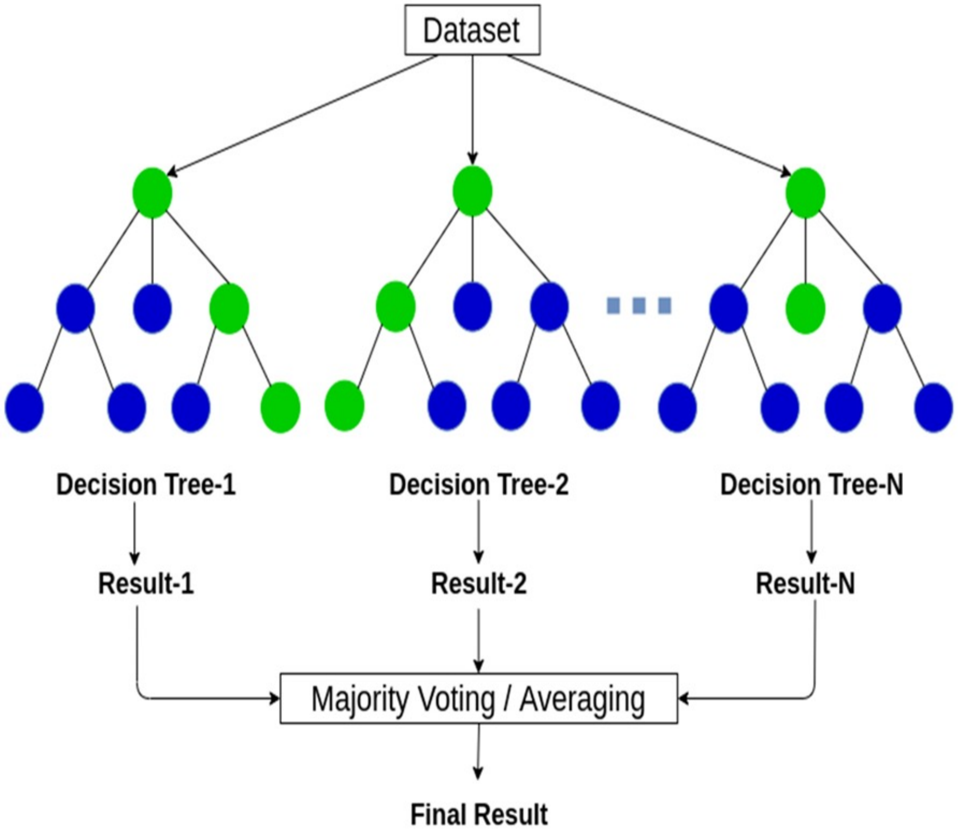
casfam - Cash benefits include family allowances and maternity leave.

kindfam- benefits in kind include early childhood education, care, and home help/accommodation.

flfp- Female Labour Force Participation Rate is calculated as the labour force divided by the total working-age female population (15-64).

magebirth- The mean age at childbearing is the mean age of mothers at the birth of their children.

Method- Random Forest (RF)



- RF is non-parametric and distribution-free machine learning approach.
- It combines the predictions of multiple decision trees.
- Decision trees are the basic elements of random forests which split the data into subsets based on the most important features.
- The split that reduces the total mean squared error (MSE) most will be chosen.
- The final prediction is the average of all individual tree predictions.

Method

RF builds an ensemble of decision trees and averages their predictions, reducing variance and avoiding overfitting.

Unlike linear models, RF can model complex interactions between variables without explicit specification.

RF can manage datasets with many features because it randomly selects subsets of features at each split and reduce correlation among trees.

Because RF predictions are aggregated across many trees, it is less sensitive to noisy data or outliers compared to more conventional approaches.

Results

Total Fertility Rate (TFR) – Model Performance

Year	Outcome	Test R^2	Test RMSE	OOB R^2	R^2 Gap
2012	TFR	0.906	0.074	0.939	-0.033
2013	TFR	0.835	0.093	0.946	-0.111
2014	TFR	0.841	0.086	0.943	-0.102
2015	TFR	0.815	0.086	0.943	-0.128
2016	TFR	0.786	0.084	0.94	-0.154
2017	TFR	0.738	0.09	0.944	-0.206
2018	TFR	0.729	0.093	0.943	-0.213
2019	TFR	0.743	0.094	0.949	-0.206
2020	TFR	0.777	0.087	0.947	-0.170
2021	TFR	0.563	0.151	0.939	-0.376
2022	TFR	0.664	0.201	0.941	-0.278

Life Expectancy (LE) – Model Performance

Year	Outcome	Test R^2	Test RMSE	OOB R^2	R^2 Gap
2012	LE	0.919	0.526	0.96	-0.041
2013	LE	0.876	0.665	0.968	-0.092
2014	LE	0.797	0.827	0.964	-0.167
2015	LE	0.885	0.616	0.965	-0.08
2016	LE	0.851	0.692	0.964	-0.113
2017	LE	0.887	0.594	0.96	-0.073
2018	LE	0.905	0.547	0.971	-0.066
2019	LE	0.856	0.671	0.968	-0.112
2020	LE	0.766	0.953	0.964	-0.198
2021	LE	0.726	1.22	0.954	-0.228
2022	LE	0.91	0.363	0.956	-0.046

Metric (paired by year)	Spearman ρ	Two-sided p-value
Test R^2: TFR vs. LE	0.09	0.79
R^2 Gap: TFR vs. LE	0.10	0.76

Variable Importance Summary

Outcome LE

Rank (RI)	Variable	Mean Importance	SD	Relative Importance (RI)	Rank (Mean)
1	emplopar	68.2	24.8	2.75	2
2	magebirth	66.7	25.4	2.63	3
3	fampatt	57.3	22.7	2.52	4
4	region	79.4	32.8	2.42	1
5	casfam	48	21	2.28	7
6	gwg	20.2	8.92	2.27	10
7	kindfam	39.9	23.8	1.68	9
8	socspen	48.5	35.1	1.38	6
9	flfp	17	12.8	1.32	11
10	edu	49	37.1	1.32	5
11	socpast	43.5	33.2	1.31	8
12	emplo	10.7	12.3	0.87	12

Outcome: TFR

Rank (RI)	Variable	Mean Importance	SD	Relative Importance (RI)	Rank (Mean)
1	fampatt	89.5	15.3	5.85	1
2	flfp	43.7	10.8	4.06	4
3	casfam	77.7	25.2	3.08	2
4	emplopar	32.2	11.8	2.73	9
5	kindfam	42.4	15.9	2.67	5
6	gwg	41.1	16	2.56	6
7	socspen	29.4	12.5	2.36	10
8	emplo	38.5	21.6	1.78	7
9	edu	36.2	20.6	1.76	8
10	region	53.6	32.3	1.66	3
11	magebirth	13.1	10.7	1.23	11
12	socpast	1.11	2.67	0.42	12

Table: Yearly Symmetry Between LE and TFR Predictor Importances (2012–2022)

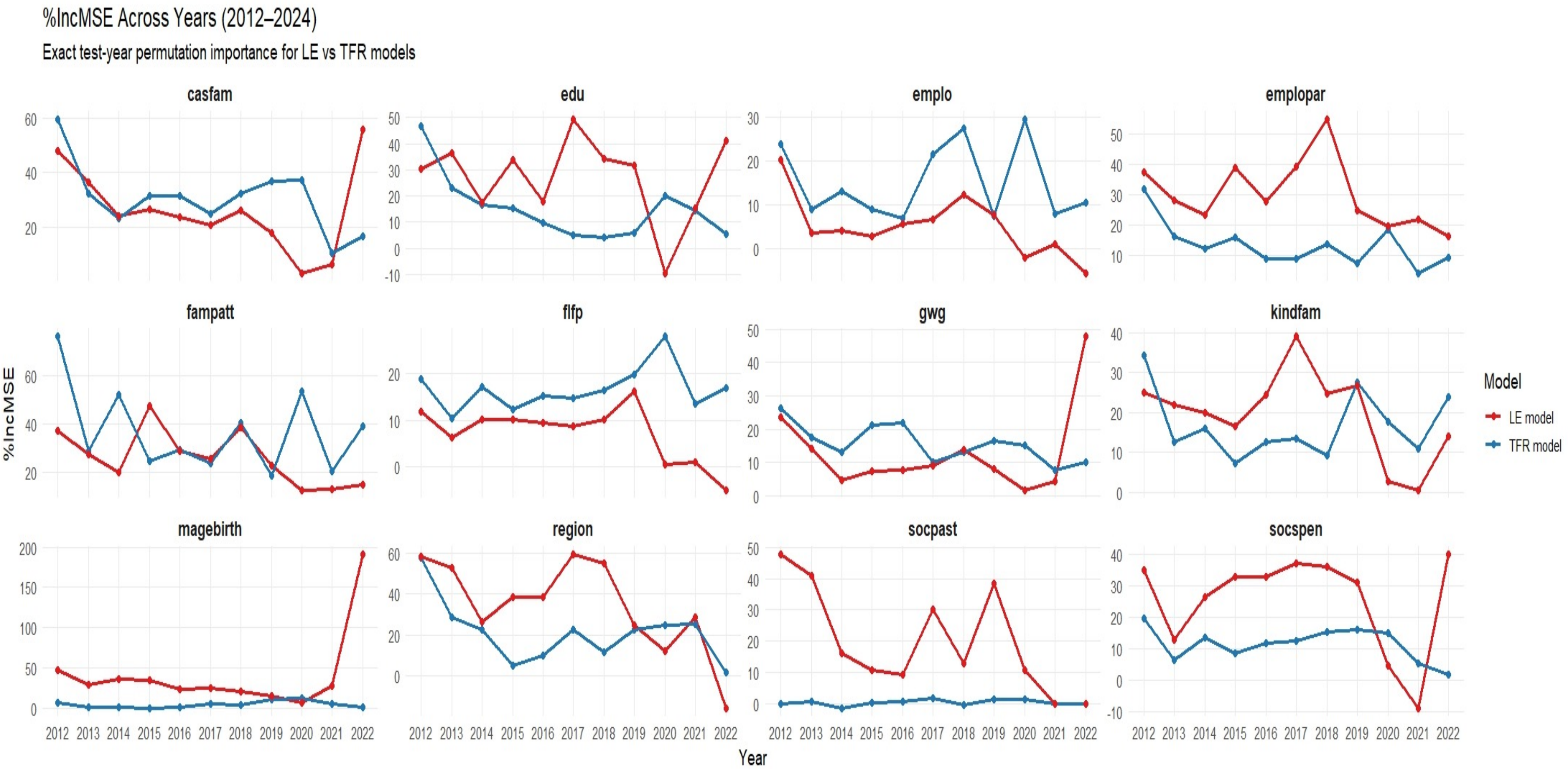
Year	Spearman ρ (LE vs TFR)
2012	0.245
2013	0.329
2014	0.084
2015	0.098
2016	0.175
2017	-0.266
2018	0.084
2019	-0.252
2020	-0.119
2021	0.392
2022	-0.105

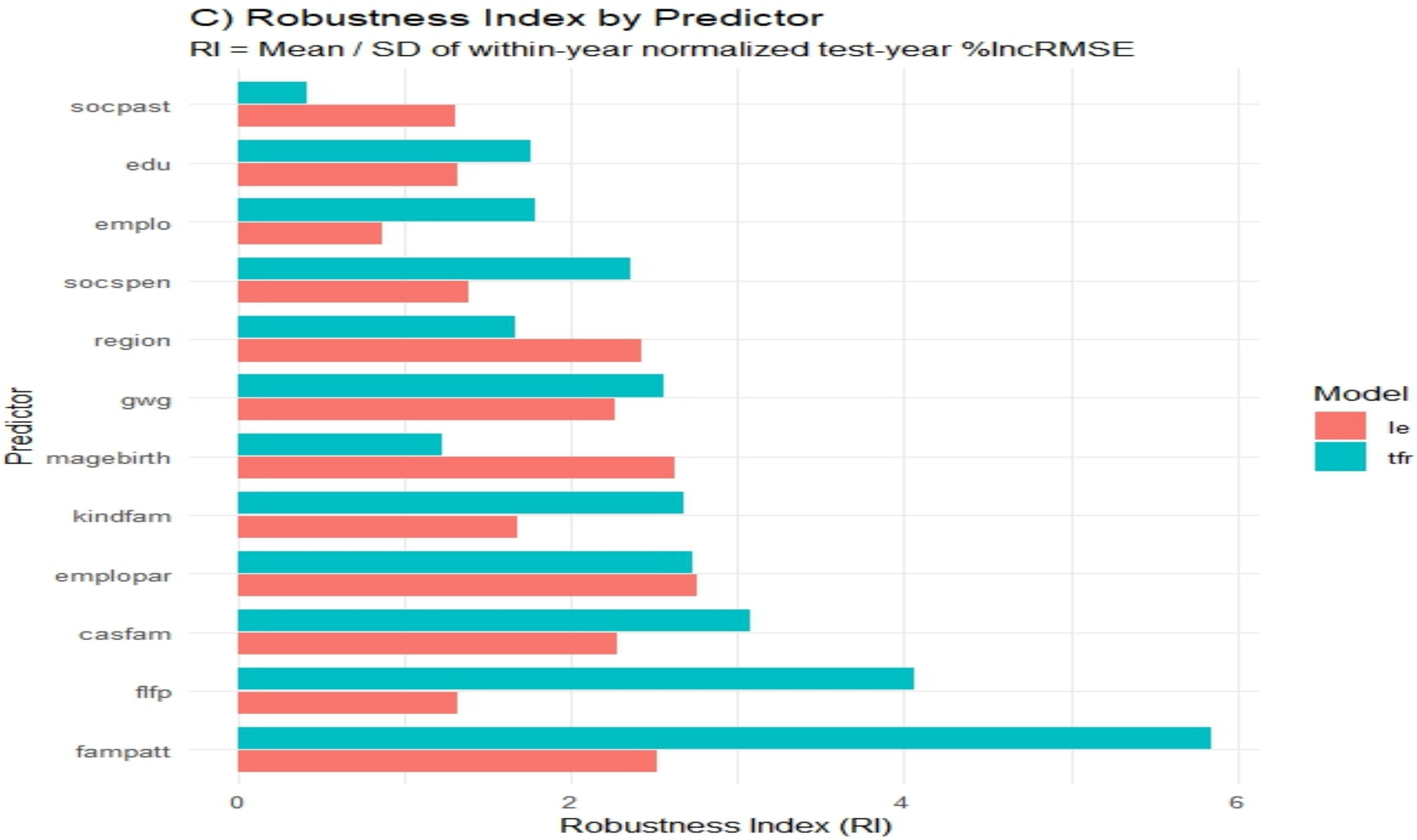
Summary Statistics

Statistic	Value
Mean ρ	0.061
Median ρ	0.084
Min ρ	-0.266 (2017)
Max ρ	0.392 (2021)
SD	0.221

Table: Test Sample Size by Year

Year	n (Test Sample)
2012	30
2013	30
2014	31
2015	31
2016	31
2017	31
2018	31
2019	31
2020	31
2021	15
2022	6


%IncMSE Across Years (2012–2024)
Exact test-year permutation importance for LE vs TFR models
casfam
edu
emplo
emplopar
fampatt
flfp
gwg
kindfam
magebirth
region
socpast
socspen
%IncMSE
Year
Model
LE model
TFR model
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022


C) Robustness Index by Predictor
RI = Mean / SD of within-year normalized test-year %IncRMSE
Predictor
socpast
edu
emplo
socspen
region
gwg
magebirth
kindfam
emplopar
casfam
flfp
fampatt
0
2
4
6
Robustness Index (RI)
Model
le
tfr

Countries with Positive Symmetry (ρ > 0.10)

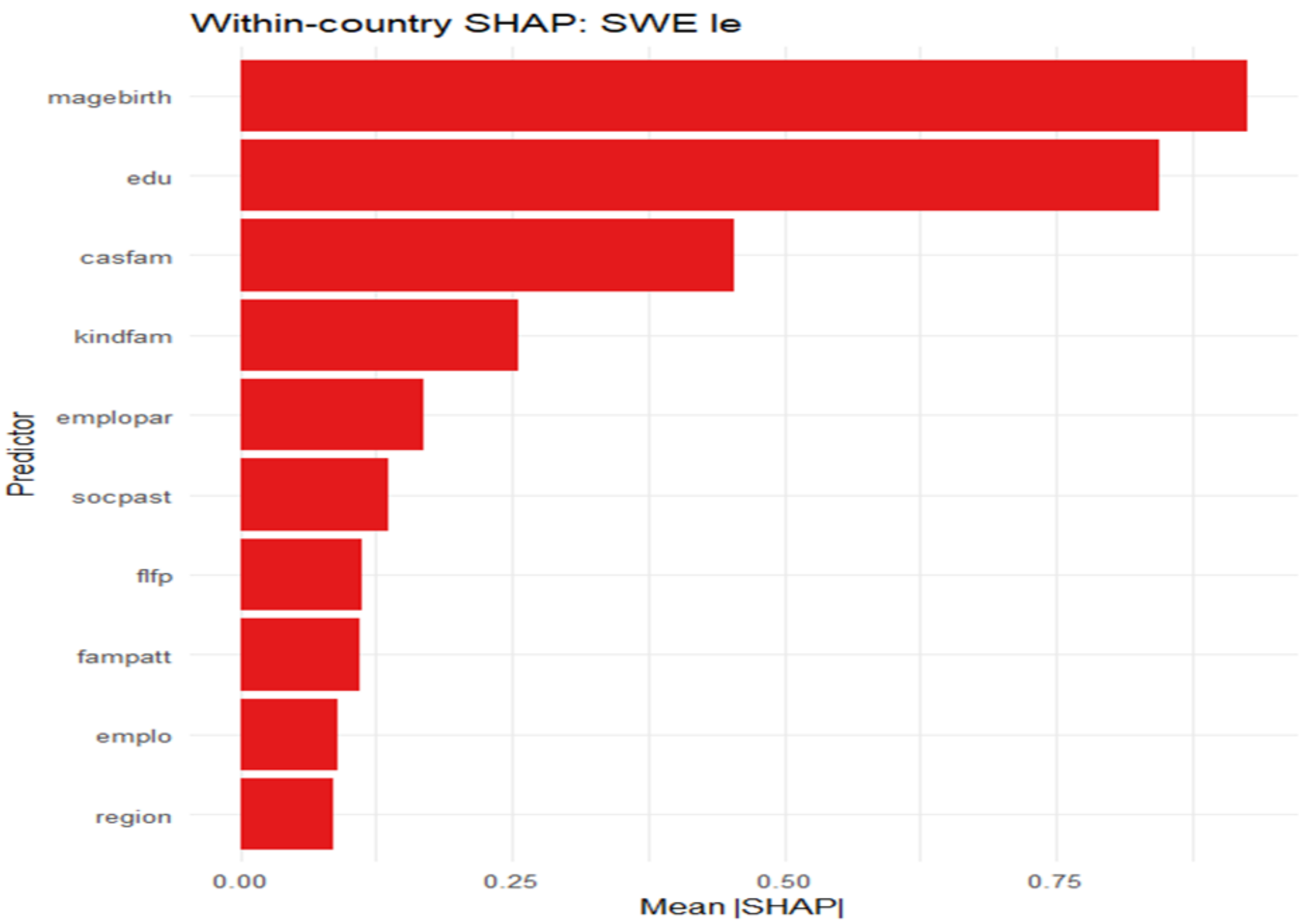
Country	Spearman ρ	n (Predictors)
Netherlands	0.741	12
Spain	0.699	12
Norway	0.552	12
Japan	0.538	12
Germany	0.503	12
Austria	0.406	12
Slovenia	0.392	12
Ireland	0.364	12
Korea	0.35	12
Belgium	0.322	12
Italy	0.287	12
France	0.259	12
Czechia	0.252	12
Denmark	0.196	12
Estonia	0.182	12
Australia	0.168	12
Finland	0.119	12

Countries with Near-Zero Symmetry (−0.10 + 0.10)

Country	Spearman ρ	n (Predictors)
Greece	0.077	12
Hungary	0.056	12
Slovak Republic	0.049	12
Iceland	0.014	12
Portugal	0.007	12
Sweden	0.007	12

Countries with Negative Symmetry (ρ < −0.10)

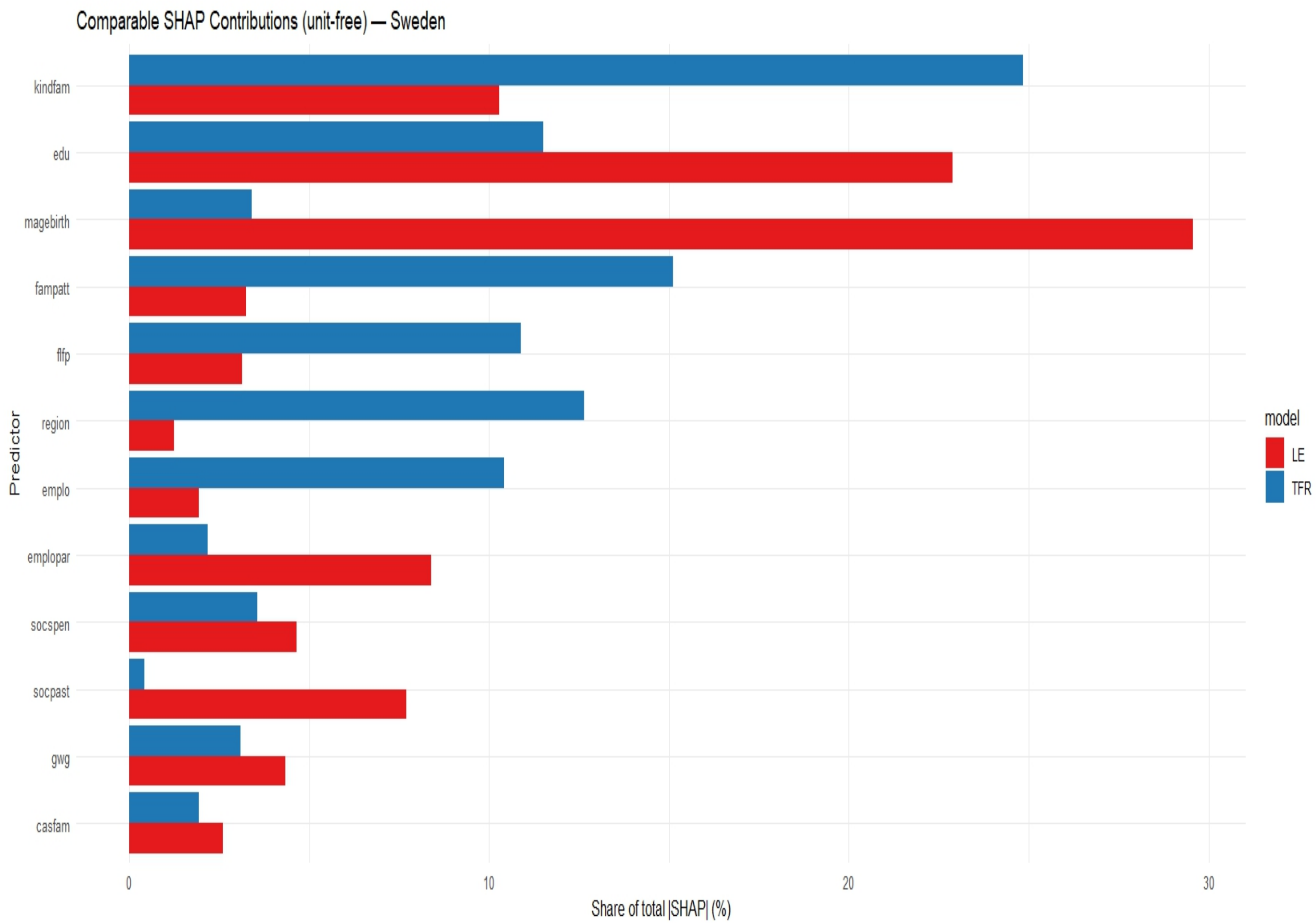
Country	Spearman ρ	n (Predictors)
Canada	-0.119	12
Luxembourg	-0.119	12
Poland	-0.154	12
Switzerland	-0.168	12
New Zealand	-0.196	12
United Kingdom	-0.252	12
Latvia	-0.483	12
Lithuania	-0.713	12



How much the variable pushes the model's prediction up or down, relative to the model's baseline?

+0.8 means the variable increases the prediction by 0.8 years while –0.8 means the variable decreases the prediction by 0.8 years

SHAP compares Sweden's real feature value (e.g., magebirth = 30 years) to many reference values sampled from the entire training dataset. It then measures how much the model's prediction changes when substituting the real value with those reference (baseline) values. This change is the SHAP value. The mean absolute SHAP is the average magnitude of these changes.



How much does each variable contribute to prediction, compared to if that variable were absent?

Limitations and Future work

Multicollinearity between predictors.

Limited Time Coverage.

Incorporate discontinuities like the Global Financial Crisis (2008–2010) and COVID-19 pandemic (2020–2021) by introducing interrupted time series.

Limited Feature Set and missing values (we need to use databases like WHO, Eurostat, and UN databases).

Train separate models for: male vs. female mortality (LE); fertility by age group or migration background; subgroup forecasting (LE) across educational groups

Divison of Ageing and Social Change (ASC)

(https://liu.se/en/organisation/liu/ikos/asc)

milos.milovanovic@liu.se

SWEAH

SWEDISH NATIONAL GRADUATE SCHOOL ON AGEING AND HEALTH

(https://liu.se/en/education/program/f7mag)

LINKÖPING UNIVERSITY

More about...

Model Specification	Mean R^2	Mean RMSE
Country-cluster	0.855	0.691
None	0.852	0.698
Time-only	**0.866**	**0.662**

mtry determines how many predictors the algorithm randomly selects when considering each split in a tree.

Small mtry : trees are more different from each and lower correlation between trees
Large mtry: trees become more similar and stronger trees but higher correlation

min_node (minimum node size) determines the minimum number of observations allowed in a terminal leaf of a tree.
It regulates tree depth and overfitting.

Small (1–3)

Trees grow very deep, model captures very fine patterns but higher risk of overfitting

Larger (10–20)

Trees stop splitting earlier, model is more generalized, lower variance but potentially higher bias

Tuning Strategy	Hyperparameter	Number of Switches
Country-cluster	min_node	7
Country-cluster	mtry	9
Country-cluster	ntrees	8
Time-only	min_node	2
Time-only	mtry	3
Time-only	ntrees	10

Tuning Strategy	Hyperparameter	Value	Count	Share (%)
Country-cluster	**min_node**	**5**	**6**	**54.5**
	min_node	8	4	36.4
	min_node	10	1	9.09
	mtry	**4**	**4**	**36.4**
	mtry	8	3	27.3
	mtry	5	2	18.2
	mtry	6	2	18.2
	ntrees	**600**	**6**	**54.5**
	ntrees	1000	3	27.3
	ntrees	800	2	18.2
None	min_node	5	26	100
	mtry	4	26	100
	ntrees	1000	26	100
Time-only	**min_node**	**5**	**10**	**90.9**
	min_node	8	1	9.09
	mtry	**8**	**8**	**72.7**
	mtry	6	2	18.2
	mtry	4	1	9.09
	ntrees	**800**	**4**	**36.4**
	ntrees	**1000**	**4**	**36.4**
	ntrees	600	3	27.3

Modeling strategy : Two evaluation loops

Time-series Tuning (inner loop)

- Training years 2000-2011
- Validation years within training are 2008-2011
- For each validation year we create expanding window: 2008 on 2000-2007; 2009 on 2000-2008, i.e., 4 folds
- For each hyperparameter combination: mtry (4,6,8,10), min_node(3,5,8),ntrees (600,800,1000) 4*3*3=36 hyperparameter sets =36 different RF setups; 36 sets * 4 folds=144 RMSE averaged across 4RMSE values 2008-2009= 36 RMSE scores, one for each hyperparameter combination.
- Hyperparameter with lowest RMSE

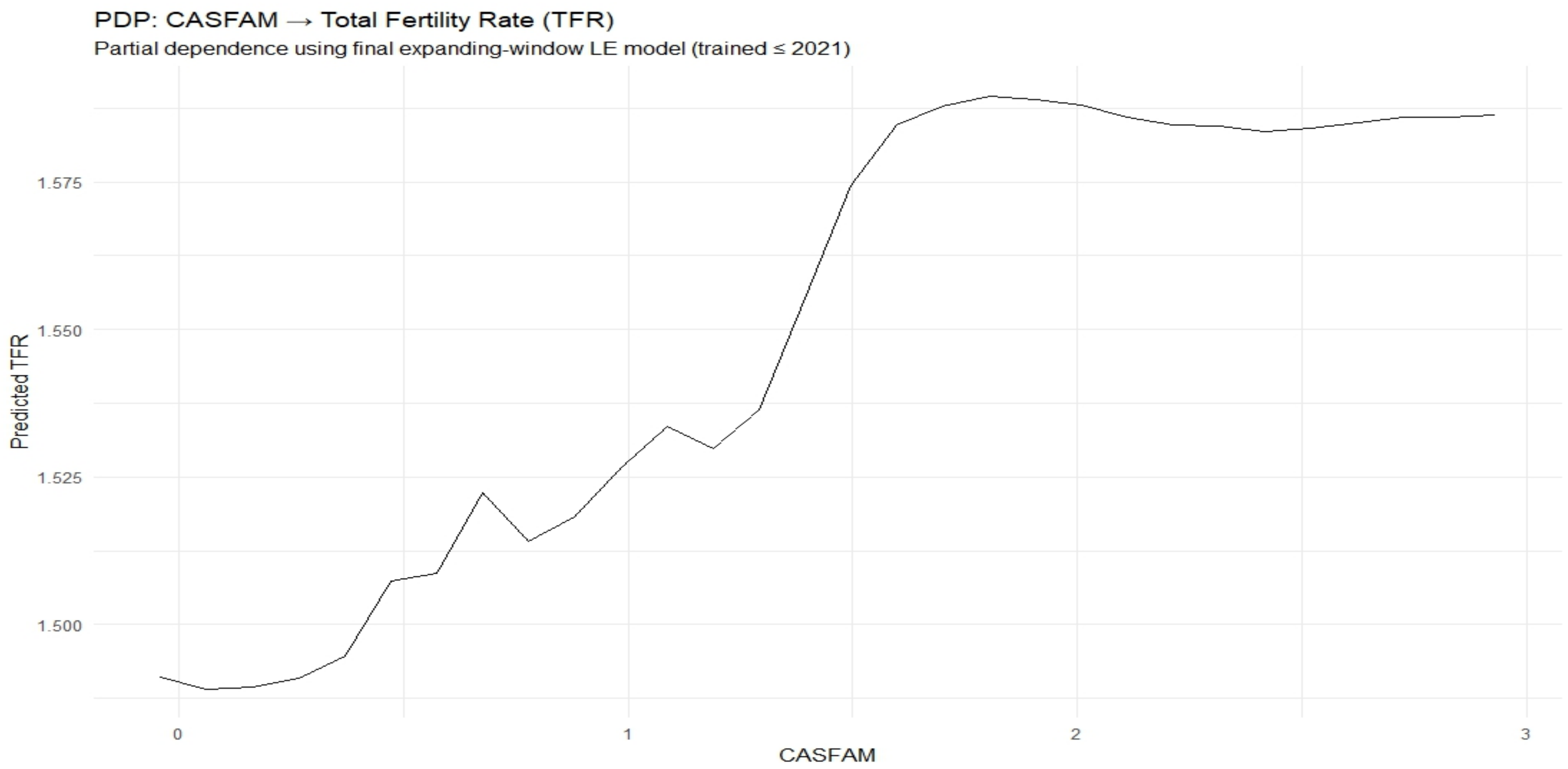
Expanding-Window Evaluation (outer loop)

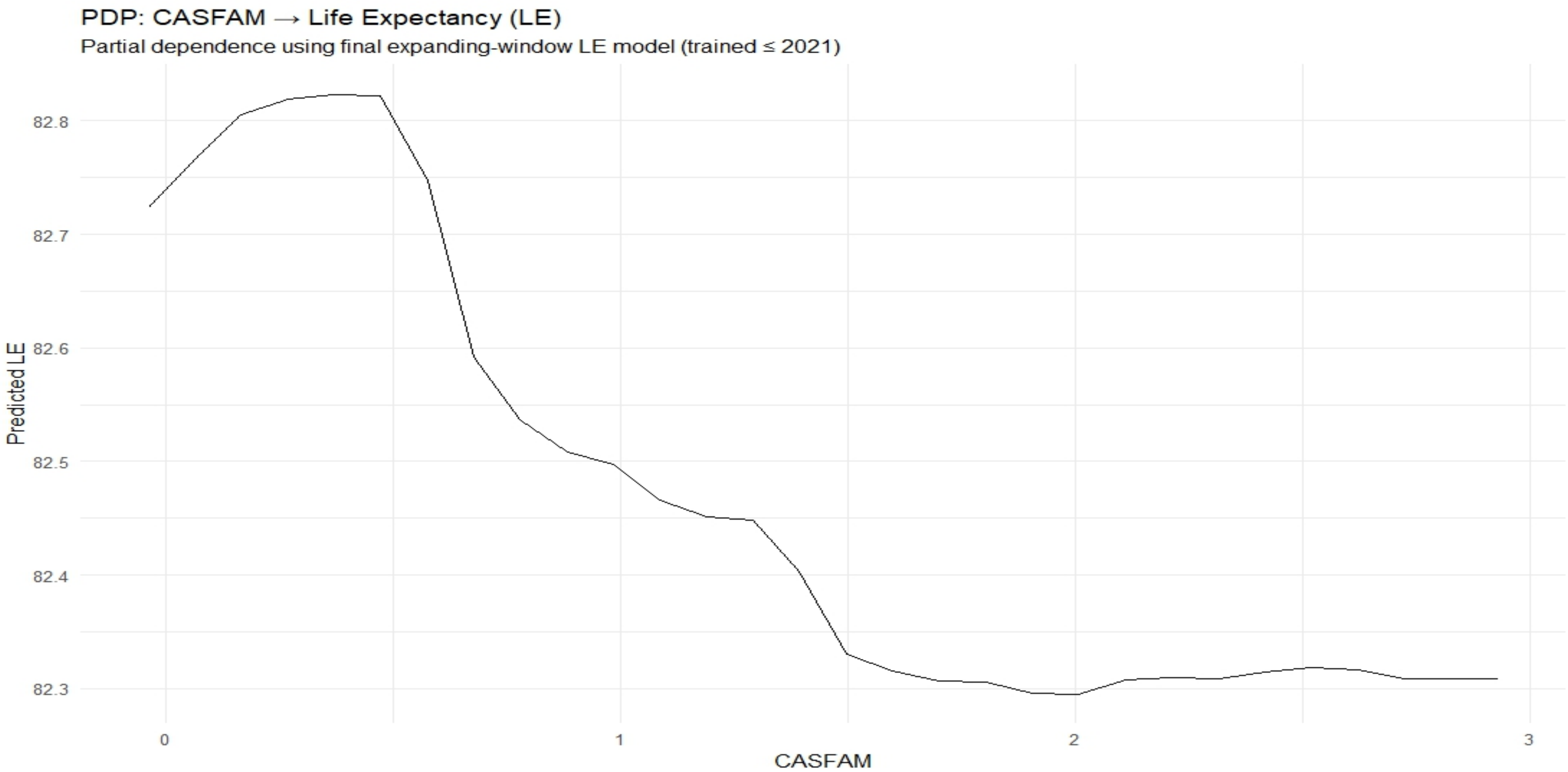
- Assess predictive accuracy over time for each validation year (2012–2024)
- Train on all data up to year-1.
- Fit the final model on the full training set using tuned hyperparameters to forecast 2012-2024
- Predict year 2012 using 2000-2011, 2013 using 2000-2012,…

"How well does the model forecast future years?"

Year	Tuning	R^2 OOB	R^2 Test	R^2 Gap
2012	Time-only	0.96	0.915	-0.045
2013	Time-only	0.968	0.871	-0.096
2014	Time-only	0.967	0.811	-0.156
2015	Time-only	0.966	0.896	-0.07
2016	Time-only	0.967	0.885	-0.082
2017	Time-only	0.969	0.928	-0.042
2018	Time-only	0.971	0.911	-0.06
2019	Time-only	0.971	0.89	-0.081
2020	Time-only	0.971	0.773	-0.198
2021	Time-only	0.961	0.74	-0.221
2022	Time-only	0.96	0.909	-0.051
2012	Country-cluster	0.959	0.912	-0.047
2013	Country-cluster	0.968	0.872	-0.096
2014	Country-cluster	0.964	0.805	-0.16
2015	Country-cluster	0.959	0.879	-0.081
2016	Country-cluster	0.96	0.848	-0.112
2017	Country-cluster	0.968	0.916	-0.052
2018	Country-cluster	0.966	0.889	-0.077
2019	Country-cluster	0.971	0.894	-0.077
2020	Country-cluster	0.965	0.765	-0.200
2021	Country-cluster	0.962	0.718	-0.244
2022	Country-cluster	0.96	0.909	-0.051
2012	None	0.96	0.915	-0.045
2013	None	0.967	0.868	-0.099
2014	None	0.967	0.791	-0.176
2015	None	0.964	0.888	-0.076
2016	None	0.965	0.845	-0.120
2017	None	0.967	0.907	-0.059
2018	None	0.968	0.891	-0.078
2019	None	0.969	0.861	-0.107
2020	None	0.968	0.77	-0.199
2021	None	0.963	0.723	-0.239
2022	None	0.958	0.917	-0.041

Year	Tuning	n_test	R^2 Test	R^2 OOB	R^2 Gap
2012	Time-only	30	0.915	0.96	-0.045
2013	Time-only	30	0.871	0.968	-0.096
2014	Time-only	31	0.811	0.967	-0.156
2015	Time-only	31	0.896	0.966	-0.07
2016	Time-only	31	0.885	0.967	-0.082
2017	Time-only	31	0.928	0.969	[illegible]
2018	Time-only	31	0.911	0.971	[illegible]
2019	Time-only	31	0.89	0.971	[illegible]
2020	Time-only	31	0.773	0.971	-0.198
2021	Time-only	15	0.74	0.961	-0.221
2022	Time-only	6	0.909	0.96	-0.051


PDP: CASFAM → Total Fertility Rate (TFR)
Partial dependence using final expanding-window LE model (trained ≤ 2021)
Predicted TFR
1.575
1.550
1.525
1.500
0
1
2
3
CASFAM


PDP: CASFAM → Life Expectancy (LE)
Partial dependence using final expanding-window LE model (trained ≤ 2021)
82.8
82.7
82.6
82.5
82.4
82.3
Predicted LE
0
1
2
3
CASFAM

Predictors

Categorical (Factors)

socpast- "1" country was member of the socialist block, otherwise "0".

fampatt-We grouped countries into clusters based on policy configurations / welfare regime types that share similar social and family policies, following Oláh, Kotowska, and Richter (2018). These clusters serve as contexts for family patterns, illustrating how demographic and gender role changes vary across different types of welfare states.

The five clusters they use are:

1. Dual-Earner (Social Democratic) Countries with strong support for work–life balance and high female employment: Denmark, Finland, Iceland, Norway, Sweden
2. Liberal Market-oriented welfare states with limited family policy support:United Kingdom, Ireland, Switzerland, USA,Canada, Australia
3. General Family Support (Conservative)Traditional systems with variable support for combining paid work and family: Austria, Belgium, France, Germany, Luxembourg, Netherlands,Japan
4. Familialistic (Mediterranean) Limited public provision; stronger gender-differentiated roles: Greece, Italy, Portugal, Spain, the Republic of Korea
5. Transition Post-Socialist Former socialist states with hybrid / diverse family policy contexts: Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia, Slovenia

Oláh, L., Richter, R., & Kotowska, I. (2018). The New Roles of Men and Women and Implications for Families and Societies. In G. Doblhammer, & J. Gumà (Eds.), A Demographic Perspective on Gender, Family and Health in Europe (pp. 41-64). Cham: Springer. https://doi.org/10.1007/978-3-319-72356-3_4

region- The region is a variable operationalized by clustering countries into five groups based o their similarity in historical development and geographical proximity. For group 1, the study included the non -European countries: USA, Canada, Japan, The Re public of Korea, and New Zealand. For group 2, the following countries were included: Austria, Belgium, France, Germany, Republic of Ireland, Luxemburg, the Netherlands, Switzerland, and Great Britain. For group 3, the transition countries (Kögel, 2004) were included, i.e., former members of the socialist block: Czecl Republic, Hungary, Poland, Slovakia, Estonia, Latvia, Lithuania, and Slovenia as the only country from the "other side of the Iron curtain". The next group are Scandinavian countries: Denmark, Finland, Iceland, Norway, and Sweden are included in group 4. Finally, group 5 includes the Mediterranean countries: Greece, Italy, Portugal, and Spain. The reason for including a variable that controls for the geographical.

socspen- According to the OECD Social Expenditure Database (SOCX), it includes government spending on: Old-age pensions, Health-related benefits, Family and child benefits, Disability benefits, Unemployment benefits, Housing support, Other social assistance programs i.e. it covers cash benefits, in-kind services, and tax breaks with a social purpose.

edu- Given the many linkages between education and family behaviour, incorporation of the share of women (aged 25 –64 years) with tertiary education as a control variable may have important consequences for study results. Since the establishment of Becker's new home economics theory (1981), the negative relationship between total fertility rates and the expansion of higher education among women is a consistent fin ding across the countries (Basten, Sobotka & Zeman,2014; Ní Bhrolcháin & Beaujouan, 2012). Recently, the results have emphasised the beneficial effect of gender equality on the increase in fertility (Impicciatore & Tomatis, 2020).

emplopar- Female part-time employment rates are defined as women in employment (whether employees or self-employed) who usually work less than 30 hours per week in their main job, anc this indicator shows the proportion of those employed part -time among all employed women (OECD, 2022c). By adding part-time employment rates, the study will be able to quantify the effec of part-time work on fertility. Arguably, in some countries (Belgium, Ireland, and The Netherlands) the use of part -time rates, in order to reconcile childbearing and work, enhance fertility (Ariza, Goiricelaya & Olazabal, 2003). Also, since infertility is associated with working extra hours, especially in young -aged workers, it can be important to control for the differences in working time (Ahn et al., 2021).

emplo- The female **employment rate**is the proportion of the working-age female population that is currently employed.

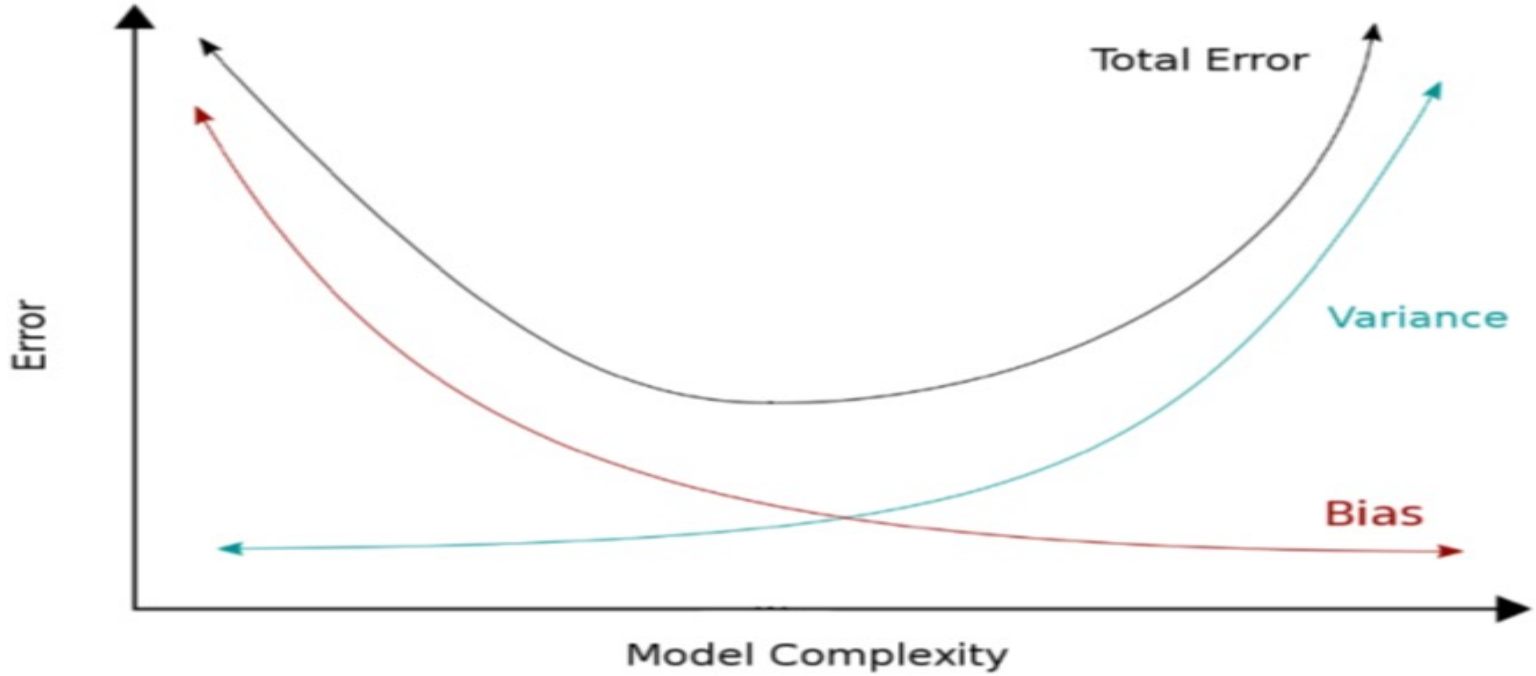
gwg- The gender wage gap is defined as the difference between the median earnings of women relative to the median earnings of men.

casfam - Cash benefits include family allowances, including maternity and parental leave, while benefits in kind include early childhood education, care, and home help/accommodation.

kindfam- Benefits-in-kind are aimed at older children and cover longer periods, thus could hav a significant effect on their success in the future (Duncan, Morris & Rodrigues, 2011; Aizer et al., 2016).

flfp- Female Labour Force Participation Rateis calculated as the labour force divided by the tota working-age female population. The working -age population refers to people aged 15 to 64 (OECD, 2022b). Since labour participation assumes females that aren't employed currently but still looking for a job, it also covered those who are not detected by unemployment statistics. Specifically, the female participation rates pool captures those who currently do not have work (also those in the army, pris on, or other institution) but actively search for a job. In contrast, by exclusively using employment rates as a dependent variable, those groups of females who are not currently employed, but are actively searching for a job would have been omitted from t he study.

Thus: the ''Bias-Variance Trade-off''



Increasing model flexibility/complexity. . .

- **Reduces** bias
- But: **increases** variance

$\Rightarrow$ A good $\hat{f}$ finds "the right balance" between the two.